Human Echolocation – Spatial Resolution and Signal Properties

Liam Norman*[1] & Lore Thaler[1]

*corresponding author liam.norman@durham.ac.uk

Department of Psychology, Durham University, Durham DH1 3LE, UK

Introduction and background

Perception is the process by which an organism uses sensory input to understand the world around them. According to some theories on perception, such as Gibson's (1979), the perceiver plays an active role in achieving this - e.g. by choosing to orient their head towards a sound source or to turn round an object in their hands. The senses that are most important for human perception of the distal environment, however, rely on some external source of energy, in the form of light or sound, to excite the environment in order to be useful. Some humans, however, have developed a mode of perception – echolocation – that is unique in that it does not require an external source of energy to reveal something about the environment. Echolocation is the perception of objects through the sound waves reflected from their surfaces, and typically an echolocator may choose to create their own sound emissions that induce these reflected sound waves. Thus, echolocation affords an individual with the unique ability to produce the energy that "excites" the world around them (Schörnich, Nagy & Wiegrebe, 2012). Given this unique characteristic of echolocation it is unsurprising that it is a technique that is used by some non-human species that have evolved to live in environments with low light levels (e.g. bats, toothed whales; Griffin, 1958; Jones, 2005), but it is quite remarkable that some blind humans have also developed this ability, often independently without any training. This is especially impressive given that the human auditory system is predisposed to actively suppress the perception of echoes – a quality often termed echo suppression or the precedence effect (Wallach, Newman, & Rosenzweig, 1949; Litovsky & Colburn, 1999). Nonetheless, some humans have become proficient at using echoes to perceive their environment and, given the unique nature of echolocation, the degree to which it can be successfully used is determined both by the acoustic properties of the sonic emission as well as the proficiency with which the returning echoes can be interpreted. This chapter will focus on the echolocation abilities of humans, describing the acoustic properties of their emissions as well as the acuity with which they are able to discriminate certain object properties – distance, angle (horizontal and vertical), size, shape and material – and the various cues that they might use to do this.

Historical accounts of blind individuals who were seemingly able to avoid obstacles and judge their distances were initially not explained by an ability to perceive echoes (e.g. see Jourdain, 1916), and instead an explanation involving "facial sense" (or "facial vision") was put forward. This explanation involved the supposition that close proximity to an object increased the pressure exerted on the individual's face. This explanation, however, was rejected following experiments by Supa, Cotzin & Dallenbach (1944), in which humans who allegedly used a "facial sense" were required to walk towards an obstacle and stop before colliding with it. When their ears were occluded, their ability to do this was significantly impaired, but not when their facial skin was covered, and other similar experiments also concluded that this recognised ability of a "facial sense" was in fact an auditory-based ability (e.g. Worchel & Dallenbach, 1947; Cotzin & Dallenbach, 1950; Köhler, 1964). These findings on humans mirrored those on the navigational abilities of bats by Donald Griffin, in which it was shown that obstructing either the bat's mouth or ears, but not their eyes, impaired their ability to navigate in flight (Griffin & Galambos, 1941). The term echolocation, coined by Griffin (1944), was thus adopted to describe navigation and obstacle avoidance by humans and nonhuman animals.

Acoustic properties of human sonar emissions

Echolocation is classically thought of as an active mode of perception, in that an echolocator typically emits a signal that excites the environment. It is also possible, however, for humans to use echolocation passively, relying on cues such as the echoes that result from external sound sources as well as the build-up of sound pressure that results from one being in close proximity to a large object (Kolarik, Cristea, Pardhan & Moore, 2014; Ashmead & Wall, 1999). Nonetheless, a very active form of echolocation is used by many blind people through the production of short sound emissions followed by the interpretation of their returning echoes. The canonical form of emission generated by human echolocators is an oral click, which is usually produced by forming a vacuum between the tongue and palate and sharply moving the tongue down (Rojas, Hermosilla, Montero & Espi, 2009; 2010), although other types of emissions have been reported to be used (e.g. finger-snapping, hand-clapping, cane-tapping and other vocalisations). The waveform of a typical oral click can be described as a sinusoidal function modulated by a decaying exponential, although the exact nature of the oral click does vary between individuals and the procedure that they use to generate it. On average, a typical click (including the exponentially decaying tail of the waveform) can last up to 15 ms (Rojas *et al*, 2009) or be as short as 3 or 5 ms, with the number of clicks made varying between 1 and 5 per second and the typical sound level varying between 60 and 108 dB (Schörnich *et al*, 2012; Thaler & Goodale, 2016). In terms of the spectral content of these oral clicks, there is energy at multiple parts of the audible range, with peak frequencies typically between 2 and 8 kHz (Schörnich *et al*, 2012; Thaler & Castillo-Serrano, 2016), which is low compared to echolocating bats, who can produce emissions in the ultrasonic range (i.e. frequencies higher than 20 kHz; Kolarik *et al*, 2014).

The oral click might be an effective form of emission for echolocation compared to other forms of natural sound emissions that are available to humans (Rojas *et al*, 2009; 2010). In comparing the use of an oral click to that of a finger-snap, for example, Tonelli, Brayda & Gori (2016) showed that novice echolocators learned just as well with either mode of emission, but at close distances to the object (30 cm) the oral click seemed to be the most useful. This oral click advantage may arise at such short ranges because the sound source of the oral click, compared to that of the finger-snap, ensures that an object directly in front of the echolocator at head-height can be effectively stimulated. Although a brief transient (i.e. click, finger snap) might be satisfactory in most natural settings, there is some evidence that acoustic emissions of a longer duration (i.e. 500 ms or longer; Schenkman & Nilsson, 2010) might also be effective (e.g. Rowan, Papadopoulos, Edwards, Holmes, Hollingdale, Evans et al, 2013; Schenkman & Nilsson, 2010; Schenkman, Nilsson and Grbic, 2016). Schenkman and colleagues (2016), for example, tested both blind and sighted participants' ability to detect when a reflecting surface was present as the number of short noise bursts made by a loudspeaker was varied between 1 and 64 per 500 ms or was a continuous noise emission. They found that at an object distance of 100cm, the blind participants' performance increased steadily with the number of clicks, with the best performance obtained with 64 bursts or with continuous noise. For sighted participants there was a decline in performance as the number of bursts exceeded 32 and, with an increase in target distance, the blind participants' performance also declined at 32 bursts. This decline was believed to result from the temporal overlap between emission and echo at the ear of the perceiver. Nonetheless, a brief oral click is accepted as the canonical form of echolocation emission, partly also because its production does not interfere with the user's breathing or movement – a rate of 2 oral clicks per second, for example, can be sustained comfortably for an indefinite period (Rojas *et al*, 2009). There is the possibility that a single type of emission is not sufficient for all instances of echolocation, and echolocators may need to adapt the type of emission they use in accordance with changing task demands, similarly to bats (Griffin, 1958) and porpoises (Tyack, 2015). There is certainly strong evidence that humans engage in head movements in order to perceive attributes such as shape or distance more accurately through echolocation (Milne *et al,* 2014; Rosenblum, Gordon & Jarquin, 2000; Wallmeier & Wiegrebe, 2014a; 2014b).

The sound emissions that are typically used by echolocators, such as oral clicks or finger snaps (and even external sounds), will reach the echolocator's ear not only through airborne vibrations, but also through bone conduction (Stenfelt & Goode, 2005). It is possible that this conduction will persist beyond the time it takes for the returning echo to reach the echolocator's ear (Patterson 1976; Schörnich et al, 2012), thus adding to the overall level of interference when perceiving the signal. The signal arriving at the ear through bone conduction, however, is predominantly low in frequency, and the interaction between this and the returning echo at different temporal delays will result in a change in the fundamental frequency of the perceived sound that Schörnich and colleagues (2012) theorised could be used as a cue by participants to judge distance to a reflecting object. The predominantly low frequency components of the emission present at the ear through bone conduction are one reason why oral clicks that contain a high level of energy in higher frequencies may be favoured over those that are lower in frequency (Schörnich et al, 2012), as the presence of high frequency sound at the ear would be a better indicator of the presence of an echo (Kuc & Kuc, 2016).

Environmental factors

In addition to the acoustics of the emission, the acoustics of the environment have been shown to affect the ability of humans to echolocate – specifically in the form of natural reverberations and additional reflectors. Although it had been thought by some that such additional factors would impair the ability to detect and discriminate objects, in a manner similar to the effect of background noise on sound localisation (Kolarik *et al*, 2014), Schenkman & Nilsson (2010) in fact showed that a reflecting object could be detected up to a distance of 100 cm in an anechoic chamber, but up to 200 cm in a conference room. Furthermore, in a study by Tonelli and colleagues (2016), in which the reverberation time was as long as 1.4 s (Tonelli *et al*, 2016), participants' ability to echolocate (both in the precision and accuracy of depth judgment) was still better in a reverberating room compared to an anechoic one. One theory for the cause of better performance in a more natural reverberating room is that "late echoes" – those echoes that reflect from walls and surrounding objects – may provide additional cues that aid in solving the task (Tonelli *et al*, 2016). Gibson (1979) posits that a surplus of perceptual information, that approximates ecological conditions, can make perception more reliable and veridical, and this principle may apply here. Nonetheless, to date these findings have only been obtained in people who were relatively new to echolocation, so it remains unclear as to whether these findings also apply to those who are more proficient.

Localising objects in depth (ranging): signal properties

The returning echo from a sonar emission such as an oral click can convey several pieces of information about the spatial content of the environment. The typical pattern of activity at the ear of an echolocator consists of the following successive sounds: the emission only, the superimposition of the emission and echo and, finally, the echo only (Kolarik *et al*, 2014), and for echoes returning from distant objects there may be a period of silence. One application of active echolocation is in the detection of an object in front of the echolocator and inferring its distance from the temporal delay between the emission and echo. Figure 1 b and d, for example, show examples of waveforms and spectrograms from binaural recordings made at the ears of an echolocator (upper panel left ear, lower panel right) when making oral clicks at a target present at a distance of 150 cm (b) or 85 cm (d). The echo shown in b has a lower amplitude to that in d and is delayed by an additional 3-4 ms. The echo returning from a reflecting object at a longer distance might be perceivable as a signal separate from the emission – and distance might be inferred from this delay - but with shorter distances the echo and emission might fuse either acoustically or perceptually. Specifically, in the case of acoustic fusion, constructive and deconstructive interference between the emission and echo take place when they are temporally superimposed, resulting in physical changes in frequency and intensity of the composite sound. In the case of perceptual fusion, illusory changes in the perception of pitch arise from the auditory system's inability to resolve small delays between the emission and the echo (a phenomenon also referred to as repetition pitch; Bilsen, 1966). The presence of an echo at close temporal proximity to the emission, therefore, may result in an acoustic and/or perceptual composite sound that has certain perceptual qualities that could still be used to infer distance to a reflector. Importantly, however, only the emission-echo delay has an unambiguous physical relationship to reflector distance, whereas the frequency and intensity of the resultant sound will also depend on size, shape and material of the reflector.

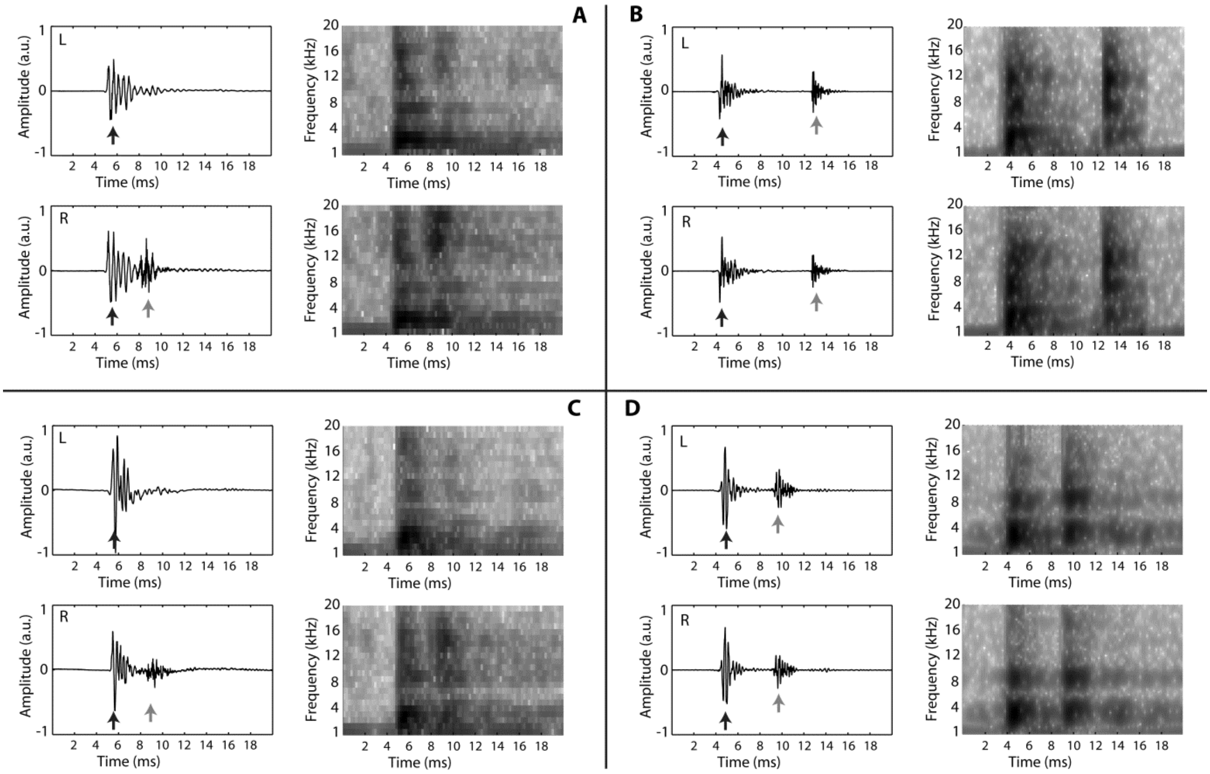

Figure 1 - Waveforms, plotting amplitude (a.u. = arbitrary units) against time (ms) and spectrograms denoting frequency (kHz) content as function of time (ms). In spectrograms darker colours indicate more energy in that frequency band at that moment in time. All figures are based on binaural recordings of clicks and echoes for four different echolocators (a–d). Recordings were made either at the entrance of the echolocators' ear canals (a, c, and d) or next to their ears, that is, on each side of the head, but placed outside the pinna, (b), while they made clicks and listened to echoes. Black arrows in waveform plots highlight clicks, and gray arrows highlight echoes. The recording sample frequency was 96 kHz for data shown on the right (b and d), and 44.1 kHz for data on the left (a and c). Spectrograms were calculated using a 1-ms window with 0.8-ms overlap in steps of 1 kHz. For (a) and (c), a sound-reflecting surface was located 60° to their right at a distance of 50 cm. For (b), a sound-reflecting surface was located straight ahead at a distance of 150 cm. For (d), a sound-reflecting surface was located straight ahead at a distance of 85 cm. Reproduced with permission from Thaler, L. , Goodale, M.A. (2016). Echolocation in People: An Overview. WIREs Cogn Sci, doi: 10.1002/wcs.1408

Schenkman and Nilsson (2011) played a set of pre-recorded sounds to both blind and sighted participants with the aim of determining which of the two cues (frequency or intensity) was more important for echolocation. The emissions in the sounds were not oral clicks, but 500 ms white noise bursts with a target object (diameter of 50 cm) placed at distances of either 100, 200 or 300 cm, with participants being required to detect the presence of the object in a 2 alternative forced choice (2AFC) task. With unprocessed sounds, blind participants detected the object almost perfectly in the 100 cm and 200 cm, but their performance dropped very close to chance in the 300 cm condition. When the sounds were processed to remove the intensity component, performance dropped by 5% but dropped by almost 20% when the frequency component was removed. This was true for both blind and sighted participants, with blind participants outperforming the sighted participants only when the frequency component was present. It is difficult to know whether these results generalise, however, as the intensity and frequency cues were not equated for their salience. Arias and Romas (1997) suggested that the level of the reflected signal was not important for discriminating the distance of the reflector but, as shown through their pitch-matching experiments of both clicks and white noise, sighted participants and one blind participant did indeed perceive echoes from closer objects as having a higher pitch than those from distant ones, and an emission-echo delay of 5 ms was optimal for evoking an illusory perception of a change in pitch (Arias and Romas, 1997; Bilsen & Ritsma, 1970). Thus, pitch might be a useful component to infer the presence of a reflector when the delay between emission and echo is short.

Localising objects in depth (ranging): spatial resolution of human echolocators

Schörnich and colleagues (2012) used a psychophysical design to measure sensitivity to depth through echolocation in sighted participants, but who were extensively trained in echolocation to a level of performance that stabilised across sessions. They used a virtual echo-acoustic space paradigm in which each participant's click was recorded by a microphone, convolved with their individual head-related transfer function and played back to them with artificially induced time delay and intensity attenuation. Head-related transfer function had been measured individually for each participant using calibrated sound sources placed at various locations in space and binaural microphones placed in a participant's ear canals. The different time delays and attenuation factors simulated sounds reflected by objects presented at different distances. They measured participants' sensitivity to changes in distance at three reference ranges: 170, 340 and 680 cm. The just-noticeable difference (JND – the minimum change in a stimulus attribute, in this case distance, required to elicit a change in perception) values at 340 cm were very consistent at approximately 80 cm across the five tested participants. The JNDs at 170 and 680 cm were more variable, but on average these were 40 and 125 cm, respectively. Two participants, however, had JNDs as high as 200 cm in the 680 cm range condition. These results can be compared to those acquired earlier by Kellogg (1962), who showed using a real setup as opposed to a virtual one, that a blind practiced echolocator was sensitive to changes in the depth of an object by 10 cm at a distance of 60 cm. Interestingly in Schörnich and colleagues' (2012) experiment, they found that the presence of a second virtual reflector at a fixed distance of 170 cm, but offset laterally at a horizontal angle (azimuth) of 45°, improved the participants' sensitivity to changes in the target reflector's distance (e.g. one participant improved from a JND of 27 cm to one of 13 cm at a reference range of 170 cm).

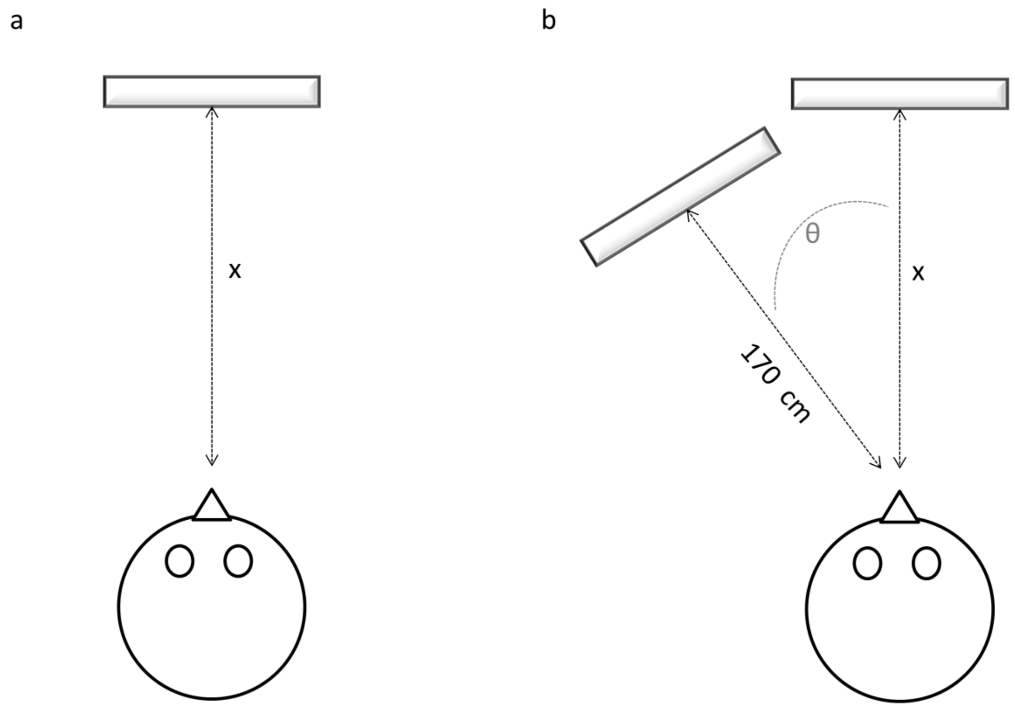

Figure 2 - Virtual stimulus setup used in Schörnich and colleagues' (2012) experiment to measure localising objects in depth (ranging). Note that reflectors used in the experiment were simulated as 'perfect acoustic mirrors', i.e. they were simulated with an artificially induced time delay and intensity attenuation. Thus, physical reflectors drawn here are for illustration only. Participants' mouth-clicks were recorded and played back to them with an artificially induced delay and intensity attenuation to simulate the presence of a reflector at a variable distance (x). Participants' sensitivity to changes in the reflector's depth was measured at three levels of depth (*x* = 170, 140 or 680 cm) either without a secondary lateral reflector (a) or with a secondary lateral reflector (b). The lateral reflector, when present, was positioned at a fixed depth of 170 cm from the participant and offset by a horizontal angle (θ) of either 15°, 30° or 45°. Participants' sensitivity to localising the primary reflector improved when the lateral reflector was present at an angle of 45°, but not 15° or 30°.

This improvement was greater with increasing azimuth of the second reflector. It was theorised that this second reflector served as a calibrating tool for the participants – a point of temporal reference against which to judge the echo from the target reflector. The second reflector did not improve performance if it was at an azimuth of 15° or 30°.

Tonelli and colleagues (2016) also trained sighted individuals without any prior experience echolocating in a depth discrimination task. Unlike the design implemented by Schörnich and colleagues (2012), that used by Tonelli and colleagues (2016) involved depth discrimination of real (i.e. non-virtual) objects in front of the echolocators. The reflector used was a sheet of poly-methyl methacrylate that was placed at a distance of either 30, 60, 90, 120 or 150 cm, with the size of the object being adjusted to keep its acoustic size constant at 10° in azimuth and 62° in elevation (see figure 3). The acoustic size of an object refers to the acoustic angle subtended by that object, and this will change for an object of fixed size as a function of its distance from the perceiver. Participants in Tonelli and colleagues' experiment (2016) had to indicate which of these 5 positions they believed the object to be at and required only one hour of training for their performance to improve. Additionally, over two training sessions their error in precision in estimating distance decreased from approximately 25 cm to approximately 15 cm when tested in a reverberant room (performance was slightly worse in an anechoic room).

Localising objects in the horizontal plane (azimuth): signal properties

Echolocation allows the user the ability to detect the position of objects in the horizontal plane, known as azimuth. Figure 1 a and c illustrate how the relative amplitudes of the echoes in the right and left ear convey information about the lateral position of the reflecting object relative to the orientation of the echolocators' head. In both cases the object was positioned at a horizontal angle (or azimuth in spherical coordinates) of 60° clockwise at a distance of 50 cm, and it is clear that the returning echo is much stronger in the right ear compared to the left. Another physical cue to lateral position is the interaural timing difference- that is, an object located to the right of the echolocator will reflect echoes to their right ear at an earlier time compared to those that reach the left ear.

Rowan and colleagues (2013; Rowan, Papadopoulos, Edwards and Allen, 2015) measured sighted participants' ability to identify whether a 55x55 cm board was presented 17° to the left or right of centre in virtual acoustic space at various distances between 60 and 180 cm. By using different bands of noise as sound emissions, they showed that the binaural cue that was used to discriminate between azimuthal differences in their study was within the frequency range of 2-12 kHz. This would highlight the importance of interaural level differences for localization. Additionally, by selectively removing binaural cues (and preventing the use of monoaural cues) they concluded that participants do indeed use a binaural cue that likely corresponds to the interaural level difference to resolve azimuthal differences. Rowan and colleagues (2015) did, however, state that one participant in their experiment was able to use a monaural low-frequency cue to discriminate between the azimuthal positions. The participant anecdotally reported hearing a change in pitch between the left and right positions, but it was not determined experimentally whether this cue was indeed being used by this participant. It is unclear if these findings will generalize to click-emissions.

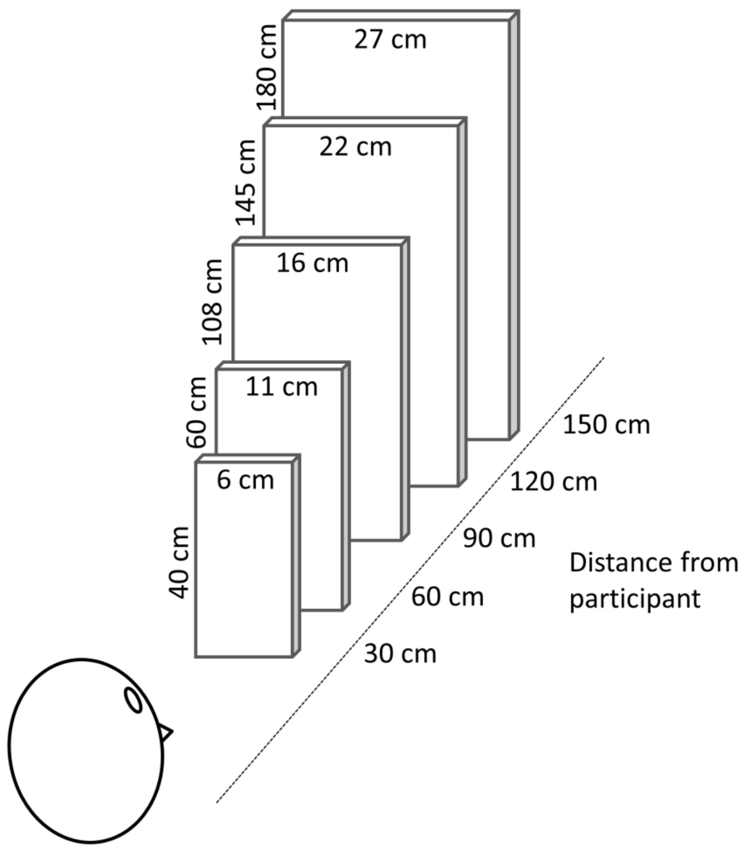

Figure 3 - Illustration of the stimuli and experimental setup used by Tonelli and colleagues (2016) to measure the resolution of localising objects in depth. Participants stood in front of a bar positioned at one of five locations in depth (30, 60, 90, 120 or 150 cm) and had to judge the depth that the bar was positioned at using echolocation with either a mouth-click or finger snap. The acoustic size of the bar (the acoustic angle it subtended) was kept constant (10° azimuth and 62° elevation) by using bars with greater height and width at greater positions in depth.

Localising objects in the horizontal plane: spatial resolution of human echolocators

Thaler, Arnott & Goodale (2011) showed using a 2AFC task that an early-blind blind echolocator could detect a change in azimuth of 4° of a pole (width 6 cm x height 180cm) placed at a distance of 150 cm. Teng, Puri and Whitney (2012) measured the sensitivity of blind echolocators to the horizontal offset of two flat circular disks (20.3 cm in diameter) at a distance of either 50 or 100 cm (see figure 4). Participants had to indicate whether the top disk was positioned to the left or right of the bottom disk. At the 50 cm distance the top disk could be horizontally offset from the bottom disk between 1.1° and 13.2° to the left or right. At the 100 cm distance it could be offset between 0.57° and 3.4° to the left or right. The three best echolocators could discriminate between angular offsets of less than 2° with 75% success, but there was considerable variability between participants. This variability was partially predicted by the age of blindness-onset – with those who lost their vision at an earlier age having better performance in this task. This level of acuity has been compared to that of sound source localisation in the frontomedial plane (Kolarik *et al*, 2014) and corresponds approximately to monocular visual acuity in the same task at a retinal eccentricity of 35° (Teng *et al*, 2012). In comparison, Wallmeier, Geßele and Wiegrebe (2013) tested sighted participants' ability to discriminate the azimuth of an object in virtual acoustic space. Using a 2AFC task and an adaptive psychophysical method, they found that on average participants could discriminate between azimuth changes of 6.7° (with a range between participants' ability of 4.8° to 9.2°) at a distance of 200 cm.

Localising objects in the vertical plane (elevation)

As for the cues used for possibly resolving object positions in the vertical plane (elevation in spherical coordinates), in general the incoming soundwaves are diffracted and reflected by the individual shape of one's pinna (Shaw, 1980) and, as a result, contain complex spectral cues that denote elevation as well as allow front and rear disambiguation (Musicant and Butler, 1984; Middlebrooks & Green, 1991). It is unclear to date to what degree this might apply within the context of echolocation, but future research should look at this issue.

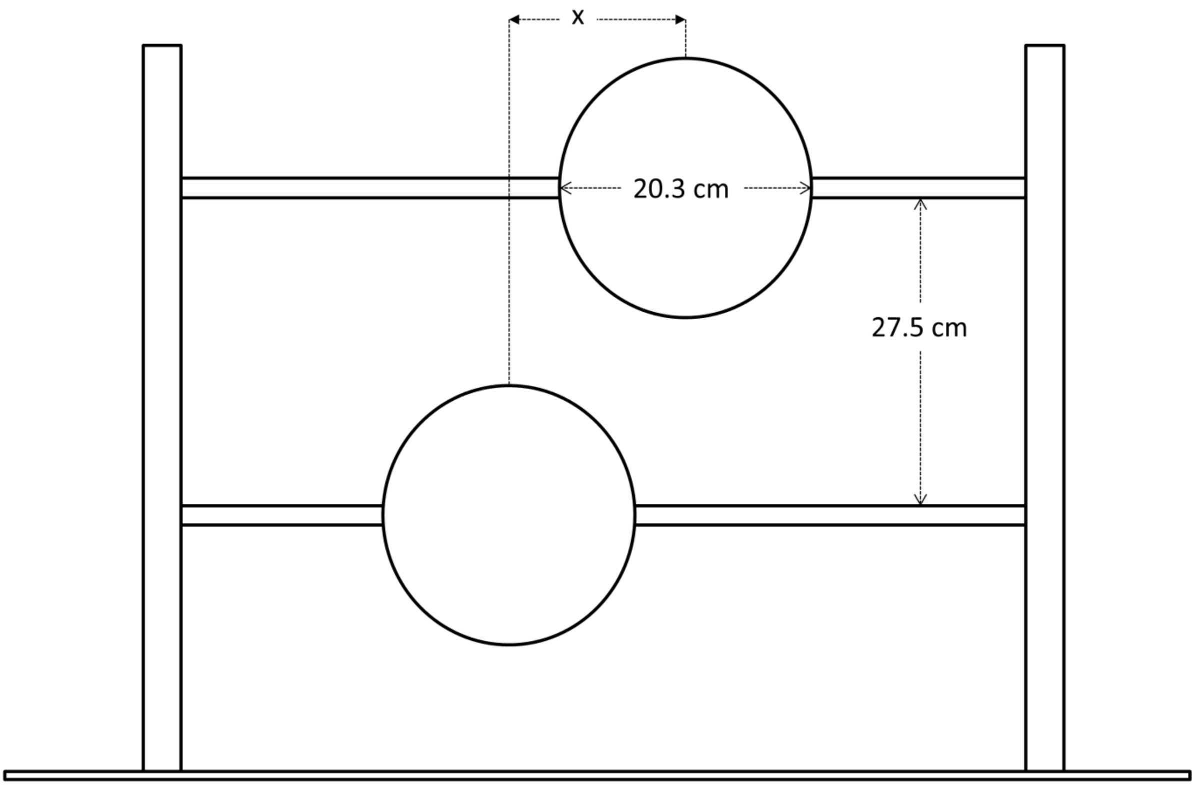

Figure 4 - Illustration of the stimulus apparatus used by Teng and colleagues (2012) to measure the spatial resolution of localising objects in the horizontal plane. Participants had to judge whether the top disk was offset to the left or right of the bottom disk. The horizontal displacement between the two disks (x) was varied such that the auditory angle subtended between the two disks varied between 1.1° and 13.2° at a distance of 50 cm from the participant or between 0.57° and 3.4° at a distance of 100 cm. The greater distance of 100 cm was used to avoid ceiling effects in performance of the three participants with the highest levels of performance.

Discriminating object size, shape and material

It is clear that human echolocators are able to resolve small changes in an object's position in space, both in terms of its depth and lateral position. It is also important, however, for human echolocators to be able to infer certain properties of an object in addition to its spatial position – such as size, shape and material – in order to appropriately interact with those objects. Teng and Whitney (2011) measured size discrimination abilities of trained sighted echolocators and compared their performance to that of an expert blind echolocator. Participants judged whether the larger of two discs was presented above or below the other, with one having a fixed diameter of 25.4 cm and the other varying between 5.1 and 22.9 cm (see figure 5). At a distance of 33 cm, the difference in the acoustic angle subtended by the two discs thus varied between 4.4° to 31.7°. Initially, participants had great difficulty discriminating even the largest differences in size, but performance improved markedly after a single training session of 100 trials. On average, performance reached an asymptote of approximately 80% for the angular size difference of 31.7°. To compare performance of the trained sighted participants to that of the expert blind participant, the highest skilled sighted participant reached a 75% performance threshold of 14.5°, whereas the same score for the expert blind participant was 8.0°. Thaler, Wilson and Gee (2014) found similar levels of performance in sighted participants with the same task, and Rice and Feinstein (1965) showed that blind participants could use echolocation to discriminate between object sizes at a surface area ratio as low as 1:1.07. Stoffregen and Pittenger (1995) state that the size of an object can be inferred through echolocation by any differences in the level or spectrum of the reflected sound that aren't accounted for by the object's distance and material. Remarkably, Milne, Anello, Goodale and Thaler (2015) have shown that a blind expert echolocator is able to infer the physical size of an object independently of its acoustic size. The expert echolocator in Milne and colleagues' (2015) study, for example, was able to correctly identify the size of a large object presented at a distance that equated its acoustic angle to that of a smaller object presented at a closer distance (see figure 6). This is known as size constancy, and is a normal property of visual perception (Boring, 1964).

The shape of an object can also be inferred using echolocation. Milne and colleagues (2014), for example, showed that expert echolocators were able to identify the shape of an object (square, triangle or two differently oriented rectangles, each with a surface area of approximately 1600 cm^2 presented at a distance of either 40 cm or 80 cm; see Figure 7) at an accuracy of 75% correct identification. Importantly, this was only possible when they were able to move their head during echolocation. Non-echolocators, however, both blind and sighted, were unable to make these discriminations at any level above chance. In contrast, Hausfeld, Power, Gorta & Harris (1982) study found that sighted participants could identify significantly above chance which of three simple geometric shapes (circle, triangle or square, with surface areas of approximately 177 cm^2) was presented at a distance of 25 cm from their forehead. The participants in that study were not practised echolocators and they performed less accurately in comparison to a blind expert echolocator, but it remains somewhat unclear as to the specific cues that are being used to solve such a task. The discrepancy between the results of Milne and colleagues (2014) and Hausfeld and colleagues (1982) might be due to the fact that performance feedback was given to those participants in the latter and/or that that objects were presented at a closer distance.

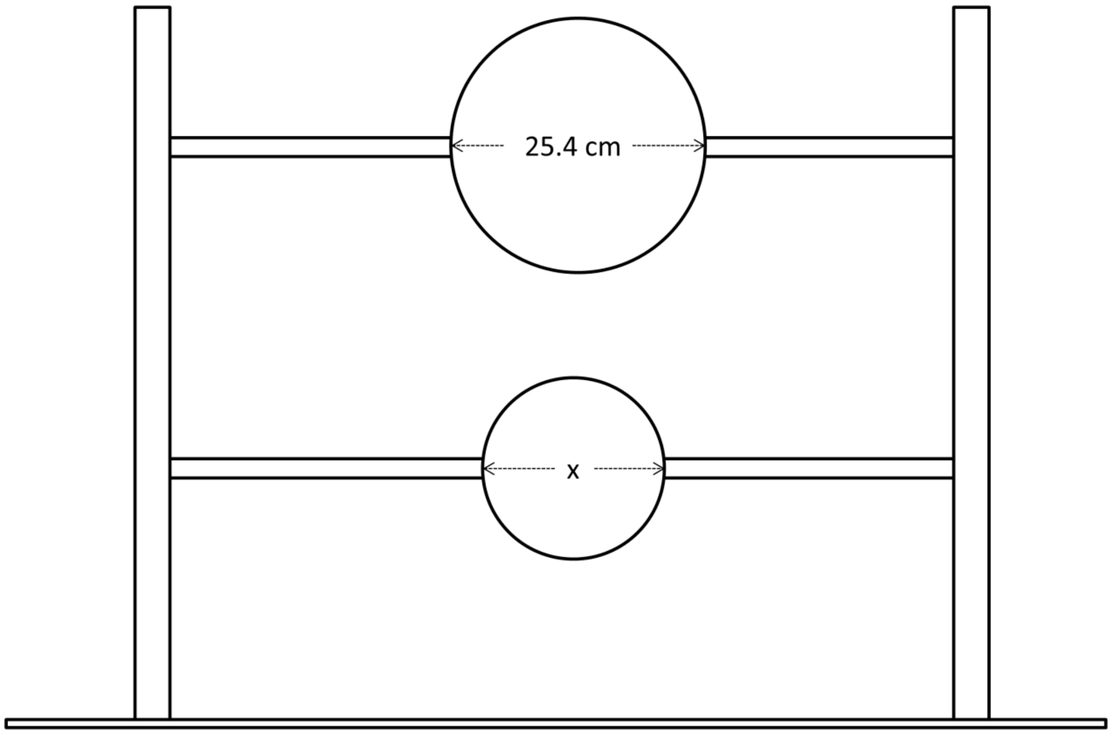

Figure 5 – Illustration of the stimulus apparatus used by Teng and Whitney (2011) to study size discrimination. The larger disk had a fixed size of 25.4 cm and would be positioned either on the upper of lower part of the apparatus, whereas the smaller disk had a variable diameter (x) between 5.1 and 22.9 cm. The auditory angle subtended by the difference in size between the two disks thus varied from 4.4° to 31.7° at a distance of 33 cm. Participants had to judge whether the larger disk was positioned on the upper or lower part of the apparatus after producing mouth-clicks.

In addition to object shape, there is also evidence that an object's material can be identified through echolocation (Hausfeld *et al,* 1982; DeLong, Au & Stamper, 2007; Milne, Goodale, Arnott, Kish & Thaler, 2005), as different materials such as carpet or wood reflect different sound frequencies in different quantities. Milne and colleagues (2014) showed that both echolocating experts as well as non-experts were able to reliably identify echoes from synthetic foliage, fleece or a whiteboard, and previous research has shown that participants report using pitch and timbre to identify materials such as wood and carpet (DeLong *et al*, 2007; Hausfeld *et al*, 1982).

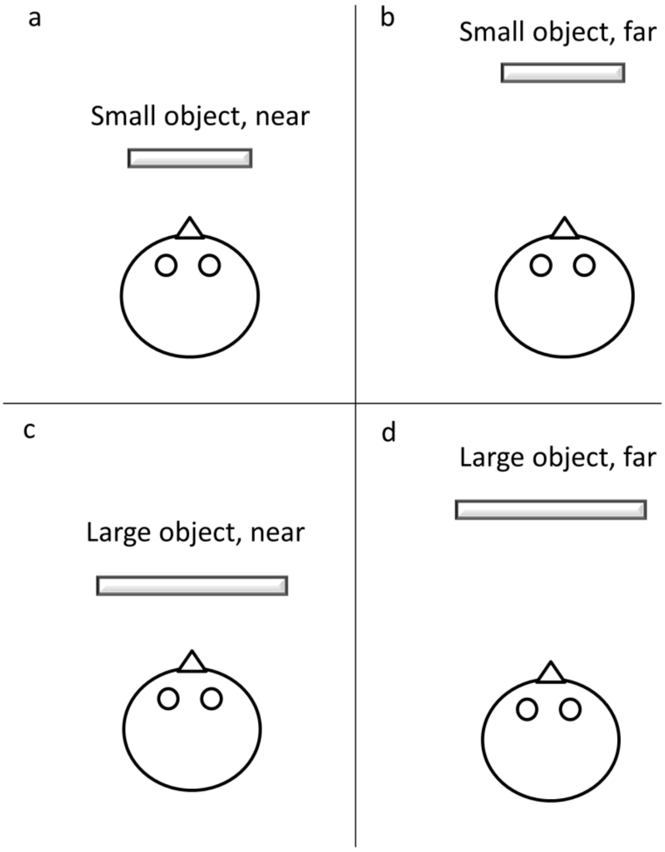

Figure 6 - Illustration of the stimuli and experimental setup used by Milne and colleagues (2015) to measure size constancy in echolocation. On a given trial, either a small or large object was positioned in front of the participant at either a near or far distance. The dimensions of the large object were such that, when positioned at the far distance (d), the object's acoustic size was identical to that of the smaller object when positioned at the near distance (a). Participants had to identify, after producing mouth-clicks, whether the small or large object was present in front of them, irrespective of its distance. Although blind and sighted novice echolocators were not able to do this, an expert echolocator was able to integrate information about acoustic size and distance in order to correctly identify the true physical size of the object, thus showing size constancy. Both rectangles and circles were used as stimuli in the experiment.

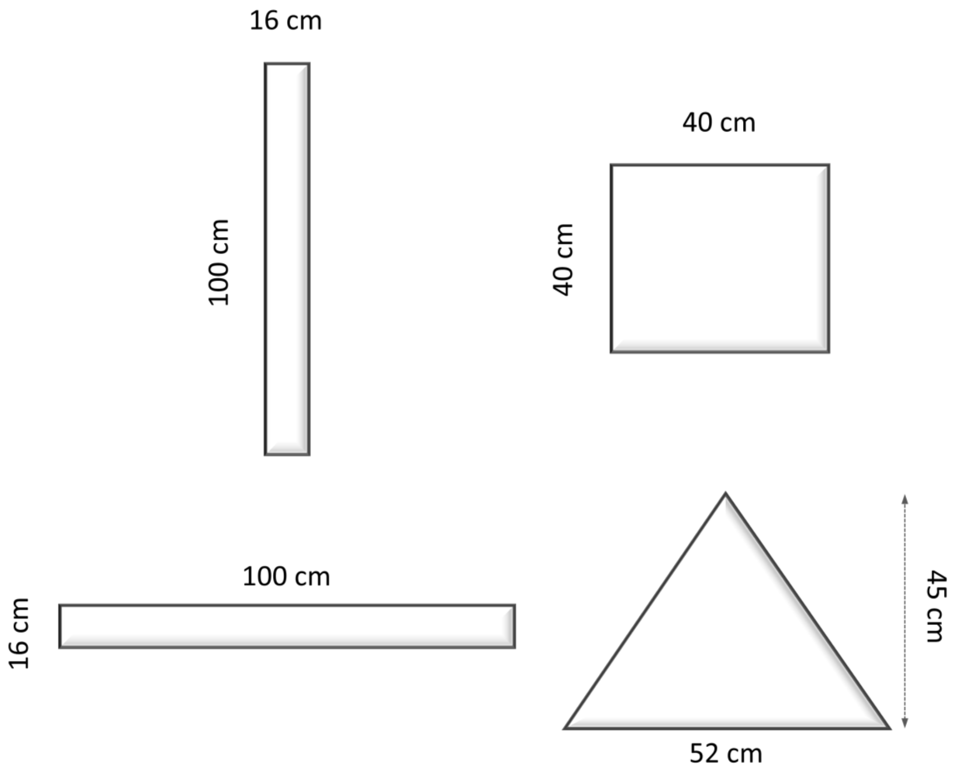

Figure 7 – illustration of the stimuli used by Milne and colleagues (2014) to study shape identification by expert echolocators and novices. The shapes were 2-dimensional and made of foam board with a surface covering of aluminium foil. The shapes were positioned on a pole with a 0.6 cm diameter, which was not detectable through echolocation, at a distance of 40 cm from the participant. Participants were presented with one of the four shapes on each trial and had to identify which shape it was. Although blind and sighted novice echolocators were not able to do this, expert echolocators were able to discriminate the shapes at a performance level of 75% correct when they were able to move their head during echolocation.

Concluding comments and future directions

To conclude, echolocation offers some humans the ability to perceive their environments without the need for an external source of energy. Humans achieve this by producing emissions, usually in the form of an oral click, and interpreting the returning echoes that are reflected from objects in the environment. Many properties of these objects can be inferred using echolocation, including distance, angular position, size, shape and material, using cues such as frequency, level, delay between emission and echo, and interaural differences in time and level of the returning echoes, and potentially other factors such as sensory-motor contingencies that are not yet fully understood. Important factors that modulate the ability to echolocate include the reverberation conditions of the surrounding environment and the presence of additional reflecting surfaces, whether the echolocator is able to move freely whilst making oral clicks and also the acoustic properties of the emissions (e.g. spectral composition, temporal duration). It is quite difficult to summarise the acuity that echolocation affords, however, given the large variability in participants' abilities, and the differences in the types of task and methods used to measure these abilities. What is clear, however, is that the acuity of some expert echolocators in discriminating certain properties such as object distance or angular position can be extraordinarily fine, and often non-expert sighted individuals can be trained to an impressive level of accuracy as well. Experiments in the future are likely to establish more common ground and likely give us a more accurate estimation of the perceptual acuity afforded by echolocation in humans and greater understanding over the factors that modulate how well echolocation can be applied.

The first scientific investigations into human echolocation were conducted in the 1940s, i.e. at about the same time when research into bat echolocation commenced. In fact, as stated in the introduction, Griffin (1944) coined the term echolocation with respect to both humans and nonhuman animals. Nonetheless, research into human echolocation has progressed at a slower pace than research into bat echolocation, for example. One possible reason for this is that there are few humans who use echolocation regularly and/or at high skill levels, making nonhuman echolocating species possibly more attractive and more available systems to investigate. In the last 5-10 years, however, research into human echolocation has experienced a resurgence. One possible reason is that humans can communicate verbally and thus provide opportunities for experimentation not available in nonhuman echolocating species. Furthermore, research has shown that some humans can achieve levels of performance that are unexpected considering that they are working in the audible (sonic) sound spectrum (e.g. Teng et al. 2012), thus highlighting opportunities for research not previously considered.

Based on the results of the experiments summarised in this chapter, it is evident that echolocation can provide humans with a range of information about the distal environment that is not limited to spatially localising an object. Specifically, the same echolocation process is used to reveal information about size, shape and material of objects as well as their spatial location. Furthermore, even though there have been only few studies to date investigating contextual factors, e.g. echolocation in anechoic environments vs. 'regular' environments, it seems that echolocation performs well across acoustic contexts (see section 'Environmental Factors'). Developers of artificial systems might therefore benefit from studies on natural human echolocation in the development of 'multi-purpose' systems that use echolocation to provide multifaceted information about the distal environment in various conditions.

It can also be learned from human echolocation that a relatively high level of spatial resolution in object localisation can be achieved with only a single sound emitter (usually the mouth) and two receivers (the ears), all working in the audible sound spectrum. Achieving good performance through echolocation in humans, therefore, is likely to be primarily limited by the neural systems processing the signals. This poses an exciting new approach to artificial sonar systems, as the focus is shifted away from sophistication of 'hardware' in creating and receiving signals towards 'software' processing and interpreting signals.

As stated before, natural human echolocation is fundamentally limited to sound information in the sonic range (<20 kHz), whereas an artificial system could go beyond these limitations to potentially achieve higher spatial acuity and directional specificity. In this context, Sohl-Dickstein and colleagues (2015) showed that an artificial system could be used in combination with a human 'processor', in that the system generates and records ultrasonic emissions and echoes and plays resultant sounds back to a human participant at a slower speed. Using such a system participants were able to successfully, and intuitively, interpret echo-acoustic cues to make judgements about an object's spatial location, suggesting that there can be a synergistic relationship between natural human echolocation and an artificial echolocation system. This presents further opportunities for research into the processes driving human echolocation and applications.

Words: 6,356 including figure captions

References

Arias, C. & Ramos, O. A. (1997) Psychoacoustic tests for the study of human echolocation ability. *Applied Acoustics, 51,* 399-419

Ashmead, D. H. & Wall, R.S. (1999). Auditory perception of walls via spectral variations in the ambient sound field. *Journal of Rehabilitation Research and Development, 36*, 313-322.

Bilsen, F. (1966). Repetition pitch: monaural interaction of a sound with the repetition of the same, but phase shifted sound. *Acustica, 17,* 295-300.

Bilsen, F. & Ritsma, R. (1969/1970). Repetition pitch and the implication for hearing theory. *Acoustica, 22, 63-73*

Boring, E. G. (1964). Size constancy in a picture. *American Journal of Psychology, 77,* 494-498.

Cotzin, M. & Dallenbach, K. (1950). "Facial vision:" the role of pitch and loudness in the perception of obstacles by the blind. *American Journal of Psychology, 63,* 485-515.

DeLong, C. M., Au, W. W. & Stamper, S. A. (2007). Echo features used by human listeners to discriminate among objects that vary in material or wall thickness: implications for echolocating dolphins. *Journal of the Acoustical Society of America, 121,* 605-617

Gibson, J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin

Griffin, D. R. (1944). Echolocation by blind men, bats and radar. *Science, 100,* 589-590.

Griffin, D. R. (1958). *Listening in the dark: the acoustic orientation of bats and men.* New Haven, CT: Yale University Press.

Griffin, D. & Galambos, R. (1941). The sensory basis of obstacle avoidance by flying bats. *Journal of Experimental Zoology, 86,* 481-506

Hausfeld, S., Power, R. P., Gorta, A. & Harris, P. (1982). Echo perception of shape and texture by sighted subjects. *Perceptual and Motor Skills, 55,* 623-632

Held, R. T., Cooper E. E., & Banks M. S. (2012). Blur and disparity are complementary cues to depth. *Current Biology,* 22, 426-431

Jones, G. (2005). Echolocation. *Current Biology, 15,* 484-488.

Jourdain, M. (1916). *Diderot's Early Philosophical Works.* Open Court, Chicago, London.

Kellogg, W. N. (1962). Sonar system of the blind. *Science, 137*, 399-404.

Köhler, I. (1964). Orientation by aural clues. *American Foundation of the Blind Research Bulletin, 4,* 14-53.

Kolarik, A. J., Cirstea, S., Pardhan, S., & Moore, B. C. (2014). A summary of research investigating echolocation abilities of blind and sighted humans. *Hearing Research, 310C,* 60–68.

Kuc, R. & Kuc, V. (2016). Modelling human echolocation of near-range targets with an audible sonar. Journal of the Acoustical Society of America, *139,* 581-587

Litovsky, R. Y., Colburn, H. S., Yost, W. A., & Guzman, S. J. (1999). The precedence effect. *The Journal of the Acoustical Society of America*, *106*, 1633-1654.

Middlebrooks, J. & Green, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology, 42,* 135-159

Milne, J. L., Anello, M. & Goodale, M. A. & Thaler, L. (2015). A blind human expert echolocator shows size constancy for objects perceived by echoes. *Neurocase, 21,* 465–470.

Milne, J. L., Goodale, M. A, & Thaler, L. (2014). The role of head movements in the discrimination of 2-D shape by blind echolocation experts. *Attention, Perception & Psychophysics, 76,* 1828–1837.

Musicant, A. D. & Butler, R. A. (1984). The influence of pinnae-based spectral cues on sound localization. *Journal of the Acoustical Society of America, 75,* 1195-1200

Nilsson, M. E. & Schenkman, B. N. (2016). Blind people are more sensitive than sighted people to binaural sound-location cues, particularly inter-aural level differences. *Hearing Research, 332, 223-232*

Palmer, S. E. (1999). *Vision science: Photons to phenomenology.* Cambridge, MA: Bradford Books/MIT Press

Patterson, R. D. (1976). Auditory filter shapes derived with noise stimuli. *Journal of the Acoustical Society of America, 59,* 640–654

Rice, C. E. & Feinstein, S. H. (1965). Sonar system of the blind: size discrimination. *Science 148,* 1107-1108

Rojas, J. A. M., Hermosilla, J. A., Montero, R.S. & Espi, P. L. L. (2009). Physical analysis of several organic signals for human echolocation: oral vacuum pulses. *Acta Acustica united with Acustica, 95,* 325-330.

Rojas, J. A. M., Hermosilla, J. A., Montero, R.S. & Espi, P. L. L. (2010). Physical analysis of several organic signals for human echolocation: hand and finger produced pulses. *Acta Acustica united with Acustica, 96,* 1069-1077.

Rosenblum, L. D., Gordon, M. S. & Jarquin, L. (2000). Echolocating distance by moving and stationary listeners. *Ecological Psychology,12,* 181–206.

Rowan, D., Papadopoulos, T., Edwards, D., Holmes, H., Hollingdale, A., Evans, L., et al. (2013). Identification of the lateral position of a virtual object based on echoes by humans. *Hearing Research, 300,* 56–65.

Rowan, D., Papadopoulos, T., Edwards, D. & Allen, R. (2015). Use of binaural and monaural cues to identify the lateral position of a virtual object using echoes. *Hearing Research, 323,* 32–39.

Schenkman, B. N. & Nilsson, M. E. (2010). Human echolocation: blind and sighted persons' ability to detect sounds recorded in the presence of a reflecting object. *Perception 39,* 483-501.

Schenkman, B. N. & Nilsson, M. E. (2011). Human echolocation: pitch versus loudness information. *Perception, 40,* 840-852.

Schörnich, S., Nagy, A. & Wiegrebe, L. (2012). Discovering your inner bat: echo-acoustic target ranging in humans. *Journal of the Association for Research in Otolaryngology, 13,* 673-682.

Shaw, E. A. G. (1980). The acoustics of the external ear. In: Studebaker GA, Hochberg I, eds: *Acoustical Factors Affecting Hearing Aid Performance and Measurement (109-125).* Baltimore: University Park Press

Sohl-Dickstein, J., Teng, T., Gaub, B., Rodgers, C., Li, C., DeWeese, M., & Harper, N. (2015). A device for human ultrasonic echolocation. *IEEE transactions on Biomedical Engineering, 62,* 1526-1534

Supa, M., Cotzin, M., & Dallenbach, K. M. (1944). "Facial vision": the perception of obstacles by the blind. *American Journal of Psychology, 57,* 133-183.

Stenfelt, S. & Goode, R. L. (2005). Transmission properties of bone conducted sound: measurements in cadaver heads. *Journal of the Acoustical Society of America, 118, 2373-2391*

Stoffregen, T. A. & Pittenger, J. B. (1995). Human echolocation as a basic form of perception and action. *Ecological Psychology, 7,* 181-216

Teng, S., Puri, A. & Whitney, D. (2012). Ultrafine spatial acuity of blind expert human echolocators. *Experimental Brain Research, 216,* 483–888.

Teng, S. & Whitney, D. (2011). The acuity of echolocation: Spatial resolution in sighted persons compared to the performance of an expert who is blind. *Journal of Visual impairment and Blindness, 105,* 20-32

Thaler, L., Arnott, S. R. & Goodale, M. A. (2011). Neural correlates of natural human echolocation in early and late blind echolocation experts. *PLoS ONE, 6,* e20162.

Thaler, L. & Castillo-Serrano, J. G. (2016). People's ability to detect objects using click-based echolocation: A direct comparison between mouth-clicks and clicks made by a loudspeaker. *PLoS ONE, 11,* e0154868.

Thaler, L. & Goodale, M. A. (2016). Echolocation in humans: an overview. *Wiley Interdisciplinary Reviews: Cognitive Science, 7,* 382-393

Thaler, L., Milne, J. L., Arnott, S., Kish, D. & Goodale, M. A. (2014). Neural correlates of motion processing through echolocation, source hearing and vision in blind echolocation experts and sighted echolocation novices. *Journal of Neurophysiology, 111,* 112–127.

Thaler, L., Wilson, R. C. & Gee, B. K. (2014). Correlation between vividness of visual imagery and echolocation ability in sighted, echo-naive people. *Experimental Brain Research, 232*, 1915–1925.

Tyack, P. (2015). Echolocation: Clicking for supper. *Elife, 4,* e07690

Tonelli, A., Brayda, L. & Gori, M. (2016). Depth Echolocation Learnt by Novice Sighted People. *PLoS ONE, 11,* e0156654.

Wallach, H. Newman, E. B. & Rosenzweig, M. R. (1949). The precedence effect in sound localization. *American Journal of Psychology. 62,* 315–336

Wallmeier L, Geßele N, Wiegrebe L.(2013). Echolocation versus suppression in humans. *Proceedings of the Royal Society B, 280,* 1769.

Wallmeier, L. & Wiegrebe, L. (2014a). Ranging in human sonar: Effects of additional early reflections and exploratory head movements. *PLoS One, 9*, e115363.

Wallmeier, L. & Wiegrebe, L. (2014b). Self-motion facilitates echo-acoustic orientation in humans. *Royal Society Open Science, 1*, 140185.

Worchel, P. & Dallenbach, K. M. (1947). “Facial vision:” perception of obstacles by the deaf-blind. *American Journal of Psychology, 60,* 502-553

Yabe, H., Tervaniemi, M., Sinkkonen, J., Huotilainen, M., Ilmoniemi R. & Näätänen, R. (1998). Temporal window of integration of auditory information in the human brain. *Psychophysiology, 35,* 615-619